

DETERMINING CUSTOMER'S WILLINGNESS TO PAY DURING IN-LAB AND REAL-LIFE VIDEO QUALITY EVALUATION

*Nicolas Staelens^a Koen Casier^a Wendy Van den Broeck^b
Brecht Vermeulen^a Piet Demeester^a*

^aGhent University - IBBT, Department of Information Technology (INTEC), Ghent, Belgium

^bVrije Universiteit Brussel - IBBT, Studies on Media, Information and Telecommunication (SMIT), Brussels, Belgium

ABSTRACT

Video has undergone rapid evolutions in the recent past, going from analogue to digital and upgrading to High Definition and 3D television. An operator will upgrade his network to provide such higher resolution and higher quality video with the aim to win as many customers as possible and/or receive a higher subscription fee from them than the competition. Clearly customers will be sensitive to higher quality, but estimating if and how much they are willing to pay extra is currently infeasible. Quality of Experience does capture realistically the customers' perception of quality, but does not automatically translate in Willingness to Pay. This paper makes statistically the relation between Quality of Experience and Willingness to Pay, based on in-lab as well as real-life quality evaluation. The outcomes of this study learn the operator how he can relate his network investments to competitive or financial advantages.

1. INTRODUCTION

Customers' perception of quality, commonly referred to as Quality of Experience (QoE) [1], is a key factor towards the success and acceptability of enhanced broadband video services such as Internet Protocol Television (IPTV) or Video on Demand (VoD) [2]. When switching to these new digital service offerings, most customers are also willing to pay in order to receive superior quality [3] compared to their traditional analogue television set.

From a service provider point of view, it is important to get an idea how the customer is experiencing the quality and, in case of a new commercial offer, how much the customer would be willing to pay for it. For existing offers, video service providers should also determine whether the offered quality remains consistent with the customer's subscription fee, not only to increase customer's satisfaction but also as a service differentiator towards competitive

offerings [4, 5].

Therefore, to estimate the effect of quality tuning in the network by means of transport protocols, server placement, over-dimensioning, etc. the operator should reason beyond traditional Quality of Service (QoS) in which he measures objective network parameters such as jitter, bandwidth and packet loss. Service providers should estimate QoE, which is a more complex indicator, depending on network parameters as well as video content type, video encoding, error concealment technique, etc. [6, 7, 8]. Even one step further, the operator should question whether the customer will be willing to pay for the offered QoE. As such, he can balance optimally his investments aimed at optimizing the quality against the expected additional revenues.

Different methods can be used to question customers about their Willingness To Pay (WTP) [9]. In the case of video quality, WTP could be measured during subjective video quality assessment [10, 11]. However, there are significant differences in quality perception when conducting video quality experiments in a controlled lab and a real-life environment [12]. This, in turn, might influence customers' WTP.

In this paper, we investigate what the monetary value is that a customer would associate to an increase/decrease in perceived quality of High Definition (HD) and Standard Definition (SD) video. More specifically, we assess the influence of visual degradations caused by network impairments (such as packet loss) on the WTP, when conducting subjective experiments within a controlled lab environment and in a real-life situation. Furthermore, we also investigate whether people are prepared to pay more to watch content in 3D.

The remainder of this paper is structured as follows. In section 2, we provide an overview of the different methods which can be used for measuring WTP and discuss some related work. The subjective experiments we conducted for measuring the WTP with respect to the offered video quality are detailed in section 3. The results of this study are presented in section 4. In Section 5, we discuss WTP for

Contact information: {nicolas.staelens,koen.casier,brecht.vermeulen,piet.demeester}@intec.ugent.be; wvdbroec@vub.ac.be.

watching 3D video content. Finally, we conclude the paper and highlight the differences between WTP in a controlled lab and a real-life environment.

2. WTP MEASUREMENT METHODOLOGIES

With the development of new digital services, questions concerning the attached value by the consumers and their WTP emerge: pay for different content types (e.g. movies, series, other TV-formats), pay for aggregation and distribution of content services, but also for content in different qualities.

There are a lot of methods in use to measure willingness to pay. Figure 1 provides an overview of different methods to measure WTP [9]. Examples are lab experiments in which consumers are asked to divide a budget over different goods and services, which is rather artificial. Field experiments in a real-life setting are less artificial. During such experiments, the focus is often on small variations in pricing in order to find the best suited pricing strategy for a certain good or service. A method often applied is the conjoint analysis, which provides the user with some prepositions on a certain good or service and the ideal pricing strategy is determined via different adaptive questions [13]. The outcome of such research is to set prices for certain services or products, but the applied methodology could be questioned as the consumer is asked to make considerations about certain goods and services that he might never consume. Furthermore these studies do not take into account specific motivations that determine the WTP, as for example the role of fandom when it concerns the consideration of paying for content or free riding [14]. The existing willingness to pay methods usually concern the measurement of the amount one wants to pay for a certain good or service (monetary indication), but do not take into account the specific dynamics and decision patterns to decide whether to pay for a certain type of content but not for another or to pay to watch a certain type of content in a better quality or not [15, 9].

Two approaches have been used in existing literature to investigate WTP with respect to the offered video quality.

First, the influence of the video bit rate of low resolution video sequences on WTP has been studied by means of a number of subjective experiments. Human observers watched a series of video sequences and were asked to indicate the amount of money they would be prepared to pay for the corresponding sequence [16] or by indicating whether they would be satisfied with the offered video quality [11]. These studies only considered encoding quality and did not take into account the occurrence of visual artifacts resulting from network impairments. The results show that there is a clear dependency between the observed video quality and WTP.

In the second approach, consumers were questioned by means of a questionnaire on their willingness to accept vi-

sual impairments in the case of discount prices for the offered video service [17, 3]. In this survey, no sequences were visually shown to the participants. The main conclusions of this study was that customers do not accept more than one visual impairment per hour and that customers are prepared to pay more for an error-free video service.

3. MEASURING WTP IN A CONTROLLED AND REAL-LIFE ENVIRONMENT

In this paper, we want to investigate the influence of network impairments on WTP by visually presenting our test subjects with a series of degraded video sequences.

In a *first experiment*, we used the standard Single Stimulus (SS) Absolute Category Rating (ACR) methodology [10] for conducting a subjective video quality experiment inside a controlled lab environment. This implies that all video sequences were presented one after another to the test subjects. Eight video sequences, varying in terms of the amount of motion and spatial details, were encoded as follows:

- *Encoder*: x264
- *Number of slices*: 1, 4 and 8
- *Number of B-pictures*: 0, 1 and 2
- *GOP size*: 15 (0 or 1 B-picture) or 16 (2 B-pictures)
- *Bit rate*: 15 Mbps

Network impairments were simulated in the H.264/AVC encoded video sequences by dropping entire NAL units which corresponds with dropping entire slices¹. Different impairments were generated by varying the type of the dropped slice, the location of this slice within the GOP and the picture, and the number of consecutive dropped slices. This resulted in a total number of 228 impaired video sequences. In order to limit the experiment duration to 20 minutes, these sequences were divided into three distinct datasets. In total, 40 non-expert viewers participated with this experiment of which some of them evaluated more than one dataset. Each dataset was evaluated by exactly 24 subjects.

After watching each sequence, subjects were required to answer the following three questions:

1. Did you perceive any visual degradations during playback?
2. How would you rate the visual quality of the sequence? (*using a 5-grade ACR scale*)
3. If you should pay for this video sequence, would you be satisfied with the video quality?

The stringent demands posed on the controlled lab environments in terms of lighting conditions, screen calibration, viewing distance, etc. are not necessarily representative for more real-life situations (e.g. watching movies in

¹An Annex B compliant encoded H.264/AVC bitstream should only consist of complete slices.

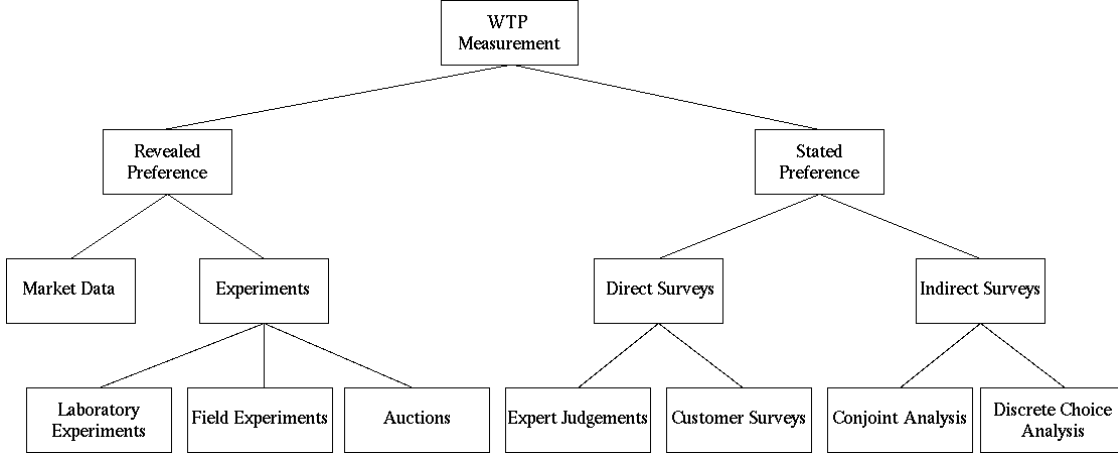


Fig. 1: Classification of the different methods for measuring WTP [9].

a living room). Furthermore, during subjective video quality evaluation, test subjects are primarily focused on quality evaluation. Therefore, in a *second experiment*, we assessed video quality in real-life environments by impairing full length DVD movies [12, 18]. Up to three impairments, equally spread over time, were injected in different movies. These impairments were either frame freezes and/or random blockiness caused by packet loss. The duration of each impairment varied between 320ms and 400ms. Subjects could watch the movie on the device of their preference (TV or computer) and were also asked to complete a questionnaire after watching the entire movie. As such, test subjects were questioned on how many impairments were perceived and their location and how they would rate the overall quality of the movie by means of a 5-grade ACR scale. For measuring the WTP, the following question was also asked to the subjects: “If you would have rented this DVD, would you mention something concerning the quality of the movie when returning it?”. 56 non-expert viewers, of which some of them watched more than one movie, participated in this experiment. In total, ratings were collected from 96 movies.

4. RESULTS

Prior to the start of a subjective experiment, test subjects receive specific instructions on how to evaluate the different video sequences. As such, they are primarily focused on visual quality evaluation. While watching television, on the other hand, people are more interested in content and are not actively evaluating visual quality. Previous research [12] already showed that this change in primary focus results in significant differences concerning impairment visibility and annoyance which, in turn, influences WTP [3].

4.1. Controlled Lab Environment

In Figure 2, the Mean Opinion Score (MOS) is plotted against the WTP for each video sequence shown during the controlled lab experiment. As can clearly be seen and as expected, there exists a strong positive linear correlation between MOS and corresponding WTP.

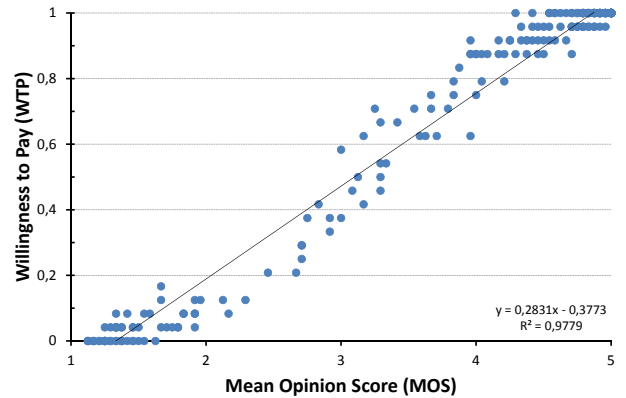


Fig. 2: Scatter plot of subjective quality ratings (MOS) against the WTP for each video sequence.

In our case, WTP can be estimated using the following equation:

$$WTP = 0.2831 \cdot MOS - 0.3773, \quad (1)$$

where MOS represents the Mean Opinion Score for the corresponding video sequence. The correlation coefficient corresponding with the above mentioned equation equals 0.98, indicating a very strong positive dependency.

In [11], WTP versus the perceived quality of small res-

olution video sequences², is approximated by: $WTP = 0.31514 \cdot MOS - 0.45664$. Applying this formula to our own obtained data also results in a correlation coefficient of 0.98. Using the Fisher Z transformation, we found that there is no statistical significant difference between the correlation coefficient obtained when applying equation 1 and the formula proposed by Ries *et al.* to estimate WTP. Hence, video resolution does not have a significant influence on WTP.

Research has shown that video content influences the perception of quality. For example, visual impairments are more rapidly perceived in high motion sequences [19]. The eight video sequences we used for this experiment can be divided into four different content classes based on the amount of motion and spatial details [10]. As video content influences perceived quality we also want to investigate whether WTP is influenced by the content type.

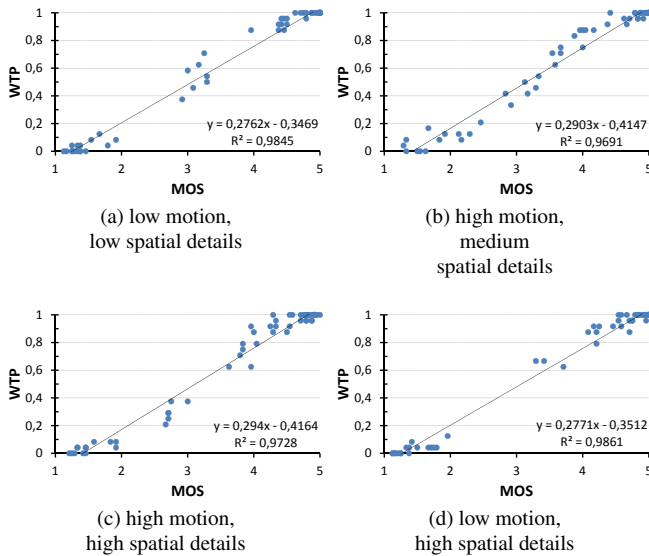


Fig. 3: Scatter plot of subjective quality ratings against the WTP, grouped by video content type.

As can be seen in Figure 3, the linear fit and correlation between WTP and MOS is only slightly content dependent. When estimating WTP per content type using equation 1, the differences between the obtained correlation coefficients are not statistically significant according to the Fisher Z transformation test. As such, in our experiment, content does not influence WTP.

4.2. Real-life Environment

During real-life QoE assessment, impairments are in general much less detected as the primary focus is on the content of the movie, not on visual quality evaluation. The

screen on which the content is watched is also of influence on the QoE, as quality expectations related to content watched on the computer are lower than those of content watched on a TV-screen. People are used to watch content of inferior quality on the computer (e.g. YouTube clips), but they expect a better image quality when watching content on the TV-set as this is what they are used to.

Concerning the WTP, during only 8 of the 96 evaluated movies (= 8.3%), the test subject indicated that he/she would say something upon returning it in case they had would have rented the movie. This was due to the fact that the perceived impairments were not found that disturbing as they were limited in frequency and length. Even when viewers do note impairments during the movie, they indicated that it is not worthwhile complaining about them as long as they do not disturb the natural fluent flow of the story. Only when the impairments disturb the natural flow of the movie and they distract viewers from being immersed in the story (e.g. when the movie stalls in an exciting scene), it ruins their viewing experience. In that case, viewers would also mention something about it if they had paid for the content. Furthermore, when impairments are perceived, people mostly assume they are caused by failing hardware (e.g. DVD-player, old television, computer too slow to process the data..) and not related to the content carrier (e.g. DVD disc) if they do not perceive physical damage like scratches.

As such, in our real-life field experiment, the majority of the test subjects are satisfied with the offered quality, even in the case when they observed up to three visual degradations. No clear relationship can be found between the Mean Opinion Score and the WTP or between the number of perceived impairments and the WTP. Based on these findings, the recommended QoE objectives as specified in ITU-T Rec. G.1080 and TR-126 can be relaxed to some extent.

5. WTP FOR WATCHING 3D VIDEO CONTENT

Nowadays, more and more movies are made available in 3D. In a recent study we questioned 21 persons, who already watched at least one 3D movie in a cinema, on their experience and their WTP for watching 3D content.

Of all respondents, 62% indicated that watching a 3D movie increases the overall viewing experience. Figure 4 provides an overview how the subjects rated their experience while watching 3D.

Watching a movie in 3D makes it easier for the viewers to immerse themselves in the movie and the story. However, only two of the respondents would watch another 3D movie. The major arguments for not watching a movie in 3D are the fact that the glasses are not comfortable enough to wear and that the higher price for watching a 3D movie does not counterbalance for the added value.

On the question whether they are willing to pay more to

²QCIF (176 x 144 pixels) and CIF (352 x 288 pixels).

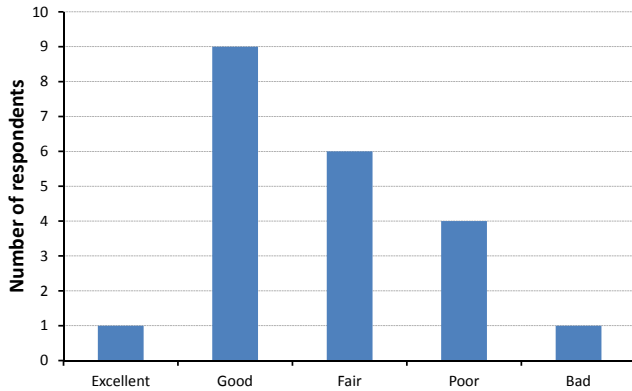


Fig. 4: Quality rating provided by the different respondents concerning their experience while watching 3D.

watch a movie in 3D, only 33% answered ‘yes’. Of all respondents who rated their experience while watching a 3D movie to be ‘excellent’ or ‘good’, 70% is willing to pay more. As already said, most of the respondents do not find 3D comfortable enough to watch. However, as display technologies keep on evolving, the current shortcoming might be resolved in the future resulting in a higher added value and viewing experience. This, in turn, might also influence viewer’s WTP.

6. CONCLUSION

In this paper, we investigated the WTP with respect to the offered video quality during a subjective experiment conducted in a controlled lab and a real-life environment. In the former, test subjects are primarily focused on visual quality evaluation whereas in the latter, subjects are more concentrated on the actual content of the video. This change in primary focus results in significant differences concerning impairment visibility and acceptance, as already pointed out in previous research.

Our results also show significant differences with respect to the WTP. When questioning subjects during a controlled lab experiment, a strong linear dependency between WTP and perceived video quality can be observed. We also found that neither video resolution nor video content have a significant influence on WTP.

During real-life QoE assessment, most of our test subjects tolerate up to three visual impairments during video playback resulting in a high WTP. Even in the event of visual impairments, most of the test subjects assumed the degradations are caused by failures of their own hardware. As long as the visual impairments do not influence immersion, viewers do not find the degradations to be disturbing and worthwhile mentioning.

We also questioned people concerning their WTP for

watching 3D movies. These results showed that this WTP is currently rather low. However, this might be due to the shortcomings of current 3D display technologies, i.e. the 3D glasses are not comfortable enough to wear.

7. ACKNOWLEDGMENT

The research activities that have been described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT) and the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT). This paper is the result of research carried out as part of the OMUS project funded by the IBBT. OMUS is being carried out by a consortium of the industrial partners: Excentis, Streamovations, Technicolor and Televis in cooperation with the IBBT research groups: IBCN, WiCa & Multimedia Lab (UGent), SMIT (VUB), PATS (UA) and COSIC (KUL).

We would also like to thank Arnaud Boussaer for conducting the survey on WTP for watching 3D video content.

8. REFERENCES

- [1] ITU-T Recommendation P.10/G.100 Amd 2, “Vocabulary for performance and quality of service,” 2008.
- [2] DSL Forum Technical Report TR-126, “Triple-play Services Quality of Experience (QoE) requirements,” DSL Forum, 2006.
- [3] G.W. Cermak, “Consumer Opinions About Frequency of Artifacts in Digital Video,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 336–343, April 2009.
- [4] A. Perkis, “Does quality impact the business model? case: Digital cinema,” in *International Workshop on Quality of Multimedia Experience (QoMEX)*, July 2009, pp. 151–156.
- [5] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin, “An overview of quality of experience measurement challenges for video applications in IP networks,” in *Wired/Wireless Internet Communications*, vol. 6074 of *Lecture Notes in Computer Science*, pp. 252–263. Springer Berlin / Heidelberg, 2010.
- [6] Y. Zhong, I. Richardson, A. Sahraie, and P. McGeorge, “Influence of task and scene content on subjective video quality,” in *Image Analysis and Recognition*, Aurlio Campilho and Mohamed Kamel, Eds., vol. 3211 of *Lecture Notes in Computer Science*, pp. 295–301. Springer Berlin / Heidelberg, 2004.

- [7] K. Wei-Ying, K. Chang-Su, and C.-C.J. Kuo, "Spatial and temporal error concealment techniques for video transmission over noisy channels," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 7, pp. 789–803, July 2006.
- [8] J. Asghar, F. Le Faucheur, and I. Hood, "Preserving video quality in IPTV networks," *IEEE Transactions on Broadcasting*, vol. 55, no. 2, pp. 386–395, June 2009.
- [9] C. Breidert, M. Hahsler, and T. Reutterer, "A review of methods for measuring willingness-to-pay," *Innovative Marketing*, vol. 2, no. 4, pp. 8–32, 2006.
- [10] ITU-T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications," International Telecommunication Union (ITU), 1999.
- [11] M. Ries, O. Nemethova, and M. Rupp, "On the willingness to pay in relation to delivered quality of mobile video streaming," in *International Conference on Consumer Electronics (ICCE 2008)*, January 2008, pp. 1–2.
- [12] N. Staelens, S. Moens, W. Van den Broeck, I. Mariën and, B. Vermeulen, P. Lambert, R. Van de Walle, and P. Demeester, "Assessing quality of experience of IPTV and Video on Demand services in real-life environments," *IEEE Transactions on Broadcasting*, vol. 56, no. 4, pp. 458–466, December 2010.
- [13] E. Brouwer and M. van der Donk, "De transitie naar digitale proposities in de krantenindustrie. onderzoek naar de betalingsbereidheid van consumenten - in opdracht van het stimuleringsfonds voor de pers," Price-WaterHouseCoopers Advisory report, 2010.
- [14] C. Buschow and A. Roth, "The relation of digital natives' fandom and their willingness to pay for music records," *Second Vienna Music Business Research Days*, June 2011.
- [15] Accent, "Review of stated preference and willingness to pay methods," Accent review report, 2010.
- [16] K. Yamori and Y. Tanaka, "Relation between willingness to pay and guaranteed minimum bandwidth in multiple-priority service," in *The 2004 Joint Conference of the 10th Asia-Pacific Conference on Communications and the 5th International Symposium on Multi-Dimensional Mobile Communications.*, August 2004, vol. 1, pp. 113 – 117.
- [17] G.W. Cermak, "Subjective Quality of Video as a Function of Frequency of Artifacts," *Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM-07)*, January 2007.
- [18] N. Staelens, B. Vermeulen, S. Moens, J.-F. Macq, P. Lambert, R. Van de Walle, and P. Demeester, "Assessing the Influence of Packet Loss and Frame Freezes on the Perceptual Quality of Full Length Movies," *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, January 2009.
- [19] N. Staelens, N. Vercammen, Y. Dhondt, B. Vermeulen, P. Lambert, R. Van de Walle, and P. Demeester, "ViQID: A no-reference bit stream-based visual quality impairment detector," in *Second International Workshop on Quality of Multimedia Experience (QoMEX)*, June 2010, pp. 206–211.